Deep Learning-Based Detection of Impacted Teeth on Panoramic Radiographs

He Zhicheng[1] (iD), Wang Yipeng[2] and Li Xiao[1]

[1]School of Computer and Information Technology, Beijing Jiaotong University, Beijing, PR China.
[2]School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, PR China.

ABSTRACT

OBJECTIVE: The aim is to detect impacted teeth in panoramic radiology by refining the pretrained MedSAM model.

STUDY DESIGN: Impacted teeth are dental issues that can cause complications and are diagnosed via radiographs. We modified SAM model for individual tooth segmentation using 1016 X-ray images. The dataset was split into training, validation, and testing sets, with a ratio of 16:3:1. We enhanced the SAM model to automatically detect impacted teeth by focusing on the tooth's center for more accurate results.

RESULTS: With 200 epochs, batch size equals to 1, and a learning rate of 0.001, random images trained the model. Results on the test set showcased performance up to an accuracy of 86.73%, F1-score of 0.5350, and IoU of 0.3652 on SAM-related models.

CONCLUSION: This study fine-tunes MedSAM for impacted tooth segmentation in X-ray images, aiding dental diagnoses. Further improvements on model accuracy and selection are essential for enhancing dental practitioners' diagnostic capabilities.

KEYWORDS: Computer-aided detection, impacted teeth, zero-shot, MedSAM

Introduction

Impacted tooth is one of the most common dental issues affecting a wide range of people of all ages. Among the most frequent services offered by oral and maxillofacial surgeons, impacted teeth have already been tentatively evaluated and managed.[1] Impacted teeth condition has prevalence rates ranging from 0.8% to 3.6%,[2] among which the third molars are commonly encountered. Radiologists typically use periapical or panoramic radiographs to diagnose an impacted tooth. While CBCT, another imaging method providing three-dimensional imaging, is also utilized in dentistry, it's not that financially friendly and convenient for impacted teeth diagnosis. Also, panoramic radiography remains widely used due to its accessibility and effectiveness in initial diagnosis, which is why this study focuses on panoramic radiographs.

Advances in computer technology have historically revolutionized radiology, enhancing patient care through standardized dental imaging diagnoses and innovative modalities. At present, given the increase in the number of patients, the emerging field of panoramic dental image inspection is facing a growing demand, while the availability of clinicians remains relatively insufficient.[3] This disparity underscores the potential of artificial intelligence (AI) in addressing this issue, catalyzing significant advancements in computer-assisted medical care. Various methodologies of computer-aided diagnosis (CAD) are actively evolving to aid in the interpretation of burgeoning volumes of medical imagery and clinical data, a topic well-documented in several published reviews.[4-6] Artificial intelligence

(AI) has played a pivotal role in advancing computer-aided healthcare, wherein neural networks, inspired by the human brain, have been instrumental.[7] Expanding the size of these networks and training them with larger datasets enable them to learn to execute specific tasks, consequently scaling up their performance for real-life applications like object detection in medical imaging.[8]

Previous studies on dental disease detection have predominantly focused on classical dental conditions such as dental caries,[9-12] periodontitis,[13-16] restorative dentistry with implants and orthodontics[17-20] While most studies have emphasized conditions like caries and periodontitis, fewer have focused specifically on impacted teeth.[21] A deep CNN system is firstly introduced to examine the location of impacted third molar teeth achieving 86.2% accuracy.[2] The following deep learning systems utilizing models such as AlexNet, VGG-16, and DetectNet were also been applied, with DetectNet achieving the highest accuracy of 96% in identifying impacted supernumerary teeth.[22] Considering from a different angle, Zhang et al[23] evaluated the predictive accuracy of an artificial neural network in post-extraction facial swelling, reaching 98% accuracy with an improved algorithm. Başaran et al[18] then utilized the R-CNN Inception v2 model to assess tooth conditions, achieving an F1-score performance of 86.25% for impacted teeth within panoramic images. To assess the potential effectiveness and accuracy of the proposed solutions on panoramic radiographs, Celik proposed a deep architecture using Faster
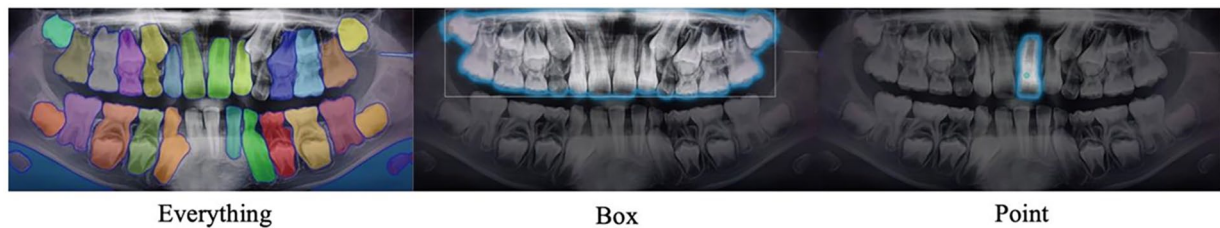
Figure 1. SAM segmentation results from different prompts.

R-CNN and YOLOv3, where YOLOv3 demonstrated superior performance (96% mAP) in detecting impacted mandibular third molar teeth.[24] These studies, though informative, are limited in number and have primarily focused on general classification and regional analyses rather than comprehensive exploration of impacted teeth.

In this paper, however, we propose to detect impacted teeth in panoramic images through pre-training and fine-tuning the Segment Anything Model (SAM)[25] on dental X-ray images. As it shows in Figure 1, the images are generated by the original SAM Model with amazing potential in segmentation, which provides fully automated segmentation with or without box and point prompts.[26] Subsequently, the center of mass from the segmentation results is extracted as a landmark for zero-shot segmentation of the tooth, aiding in identifying inconspicuous positions of impacted teeth to assist clinicians. The SAM family models[25,27] are then utilized for zero-shot segmentation of individual lesions.

We aim to push the boundaries of dental imaging research by utilizing zero-shot segmentation, a cutting-edge technique that enables accurate target detection without the need for specific training on dental data. By employing zero-shot segmentation, it significantly enhance diagnostic efficiency, offering a novel solution that could reduce the time and resources required for data annotation. To support this groundbreaking technique, we utilize a comprehensive dataset that goes beyond the scope of most existing literature. Additionally, we refine the SAM model through specialized pre-training and fine-tuning on dental X-ray images, further expanding the applicability of this model in the field of dentistry. By integrating zero-shot segmentation with these advanced methodologies, our research seeks to set a new standard in dental diagnostics, paving the way for future innovations in medical imaging.

Material and Methods

Dataset and pre-processing

The dataset[29] is provided by DENTEX, a competition held by MICCAI 2023. It is made up of 3 versions of annotations. One with quadrant information, another adds up with enumeration data and the last one with disease labels. The disease includes 4 specific categories, where the impacted teeth class has been filtrated for further training. By analyzing the json file containing labels about the segmentation mask position, the coordinates of each point are linked together to form the segmentation result. The enclosed area's RGB value was set as (000) and the rest of the picture was (255 255 255), thus the mask dataset was generated.

To enhance the performance of the training results, 254 pictures of them all have been randomly added noise, rotated, flipped, inversed color and changed lightness. As a result, 1016 samples were trained.

Model architecture

In the initial phase of our study, we employ the Segment Anything Model (SAM) to segment impacted teeth within our dataset. It is observed that SAM's zero-shot learning capabilities are suboptimal, specifically failing to discriminate between impacted and healthy teeth. We compare SAM's output to annotated ground truth data, finding that SAM's predictions frequently diverge from expert-annotated labels, particularly in distinguishing between impacted and healthy teeth, which highlights the model's limitations. As we know, the original dataset utilized for SAM does not encompass samples of impacted teeth. Furthermore, even the MedSAM, which is a SAM variant that has been adjusted for medical applications, does not include a training set for oral sections. Consequently, to enhance the precision of impacted tooth segmentation, we fine-tune the SAM utilizing methodologies delineated in MedSAM[28] on our dataset. In MedSAM, it utilizes an interactive training approach where prompts are derived from the masks for extensive prompt-based learning. After that, it applies masked autoencoding (MAE) as the objective for task-neutral self-supervised learning. We first train the SAM model using 1016 labeled panoramic X-ray images of impacted teeth, enabling it to predict possible regions of impacted tooth. However, despite obtaining the relative positions of every impacted tooth, it does not result in complete segmentation of each. Instead, it generates masks around each impacted tooth. Therefore, we take center of gravity of masks as cue points and feed them into original SAM family models, including SAM-Med2D, SAM-base, SAM-huge, SAM-large, and MedSAM. As a result, it segments the entire tooth accurately and intuitively. The whole structure and workflow can be seen below.

As demonstrated in Figure 2, the entire structure primarily consists of 2 SAM (Segment Anything) models and a MPC (Mass Prediction Component). Firstly, the input interface accepts a dental X-ray image, which is then fed into a fine-tuned SAM to output a mask indicating the relative positions of each
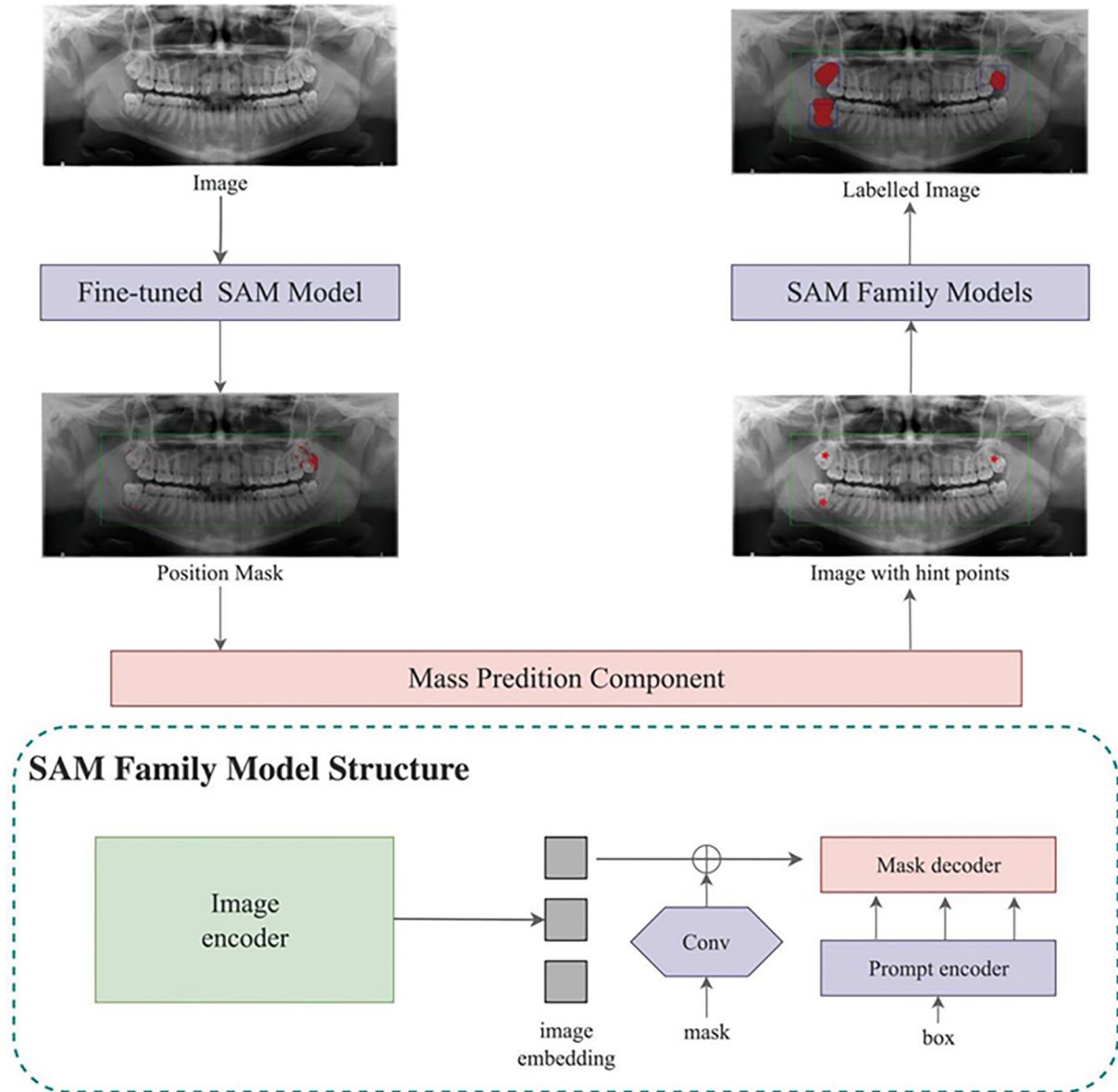
Figure 2. The workflow of our proposed structure as well as SAM family model structure.

impacted tooth. Then the mask is relayed to the MPC. The MPC is a component that generates centroids for each mask section, serving as guiding points inputted into an original, unmodified SAM model. Finally, this process results in an X-ray image comprehensively annotated with each impacted tooth.

The proposed model demonstrates the capacity to discern the presence of impacted teeth in patients based on their X-ray images. Moreover, the model exhibits a heightened precision in delineating potential impactions through a one-click prediction process. This capability allows for accurate identification and localization of each prospective impacted tooth, which contributes to an efficient diagnostic workflow in dental imaging analysis.

Results
After the fine-tuned SAM model recognizes the impacted tooth, the mask will take the center of gravity as the mark of the impacted tooth. During the training period, SAM's parameters were manually set with an epoch size of 200, a batch size of 1, and a learning rate of 0.001. The dataset contains, 1016 X-ray images, of which the author randomly selected 15% for validation. To assess the accuracy of the results, the rest 5% of the dataset were only used in testing of the network.

In this study, we evaluate the proposed structure with F1-score metrics, intersection over union(IoU), recall, and accuracy, respectively. Given that accuracy itself may not be very comprehensive, the harmonic average of precision as F1-score and sensitivity as recall are assessed. IoU method is another formulation to calculate the rate of overlap between the ground truth (masked area) and the predicted result. To gain the amount of these metrics, we need to introduce TP (true positive), TN (true negative), FP (false positive), and FN (false negative) to help with the definition. The detailed information are given as follows:

Precision = TP/(TP + FP)	(1)
Accuracy = (TP + TN)/(TP + FP + TN + FN)	(2)
Recall = TP/(TP + FN)	(3)
F1-score = 2× precision × recall/(precision + recall)	(4)
IoU = TP/(TP + FP + FN)	(5)

Table 1. The evaluation metrics for the test set's performance.

MODEL	F1-SCORE	IOU	RECALL	ACCURACY
SAM-Med2D	0.2590	0.1488	0.2064	0.3477
SAM-base	0.2806	0.1632	0.2266	0.3684
SAM-huge	0.2904	0.1698	0.2330	0.3852
SAM-large	0.3951	0.2462	0.3108	0.5421
MedSAM	**0.5350**	**0.3652**	**0.3868**	**0.8673**

Table 1 demonstrates the performance evaluation metrics for the proposed method, where MedSAM peaked in all indices. Figure 3 visualizes the performance of each model in SAM family. Among them all, MedSAM has obtained a 86.73% accuracy score. The performance of other models were not quite appealing, due to the underperformed generality of zero-shot training. However, it has been noticed that multiple results of impacted teeth were recognized by the model, which might match with the potential results. Unfortunately, the masks of the training and testing set were only annotated the most prominent impacted one tooth per image, the segmentation capability of multiple teeth needs to be tested on a more refined model.

Figure 4. shows the truncated and amplified image and mask randomly selected from the testing set, along with the output results of the Huge and Large models from the SAM family. As the compare between the mask and the obtained output, it can be seen that the produced result is highly similar to the original one.

Discussion

This study explored the application of the Segment Anything Model (SAM) for impacted tooth segmentation by fine-tuning it using MedSAM, particularly leveraging the centroid extraction approach. The fine-tuning of MedSAM with X-ray images and markers significantly improved the detection accuracy of impacted teeth in panoramic radiographs, achieving an accuracy rate of 86.73%. This confirms the model's potential to enhance dental diagnostics by accurately segmenting even subtle cases of impacted teeth. Additionally, the study demonstrated the feasibility of zero-shot learning in this context, further showcasing its efficiency.

Compared to traditional large-scale training models, the small-scale training approach used in this study yielded promising results, highlighting the adaptability of MedSAM to specialized medical datasets. Given the limited size and specificity of dental images, the ability to train effectively on a smaller dataset is a key advantage. These findings align with previous research, affirming that fine-tuning pre-trained models can significantly enhance diagnostic performance, particularly in niche areas like impacted tooth detection.

However, the study has certain limitations. The relatively small dataset may impact the model's generalizability, and the training set may not fully cover all impacted teeth, potentially affecting segmentation accuracy. Additionally, the demographic scope of the dataset may not be fully representative, which should be considered in future applications.

Further research is needed to validate MedSAM's efficacy on larger and more diverse datasets and explore its integration with other advanced diagnostic tools. Future work should also aim to refine parameter tuning and expand the model's segmentation capabilities across various dental conditions, ensuring its applicability in clinical settings. Additionally, exploring other SAM family models could help achieve better segmentation performance, further enhancing the accuracy of dental imaging diagnostics and minimizing potential misdiagnoses.

Conclusion

The current study leverages fine-tuning of the SAM model to enhance the segmentation of teeth in panoramic X-ray images, particularly for detecting impacted teeth. This research highlights the potential of using advanced deep learning techniques to improve diagnostic accuracy and efficiency in dental imaging. By accurately segmenting impacted teeth, the model contributes to better clinical decision-making, enabling more precise treatment planning. Furthermore, the fine-tuned SAM model's ability to identify subtle features in panoramic images demonstrates its applicability in diagnosing various dental conditions. This study also underscores the importance of integrating machine learning models in clinical practice, offering promising directions for future improvements in dental diagnostics.
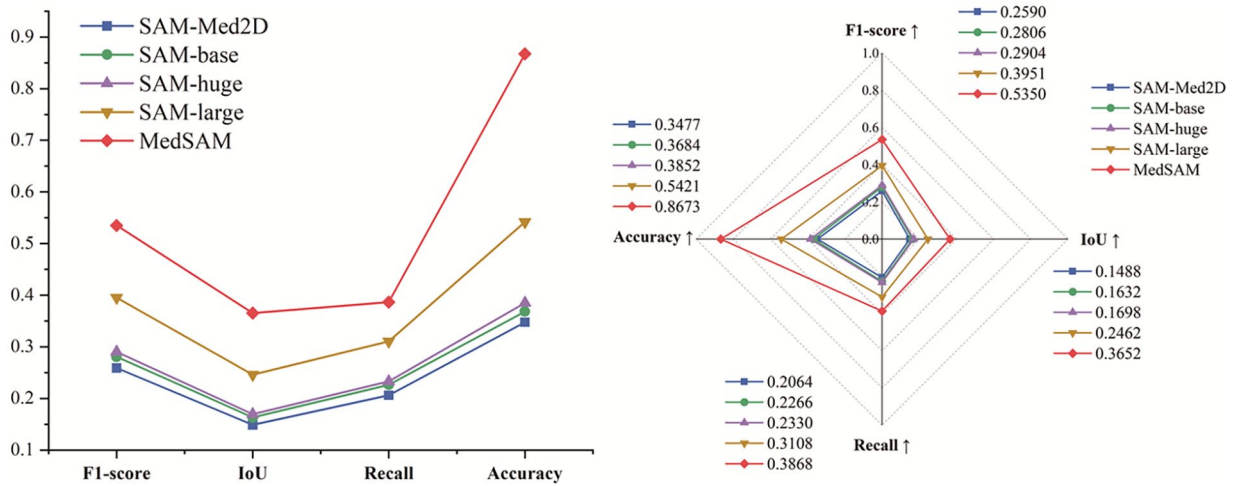
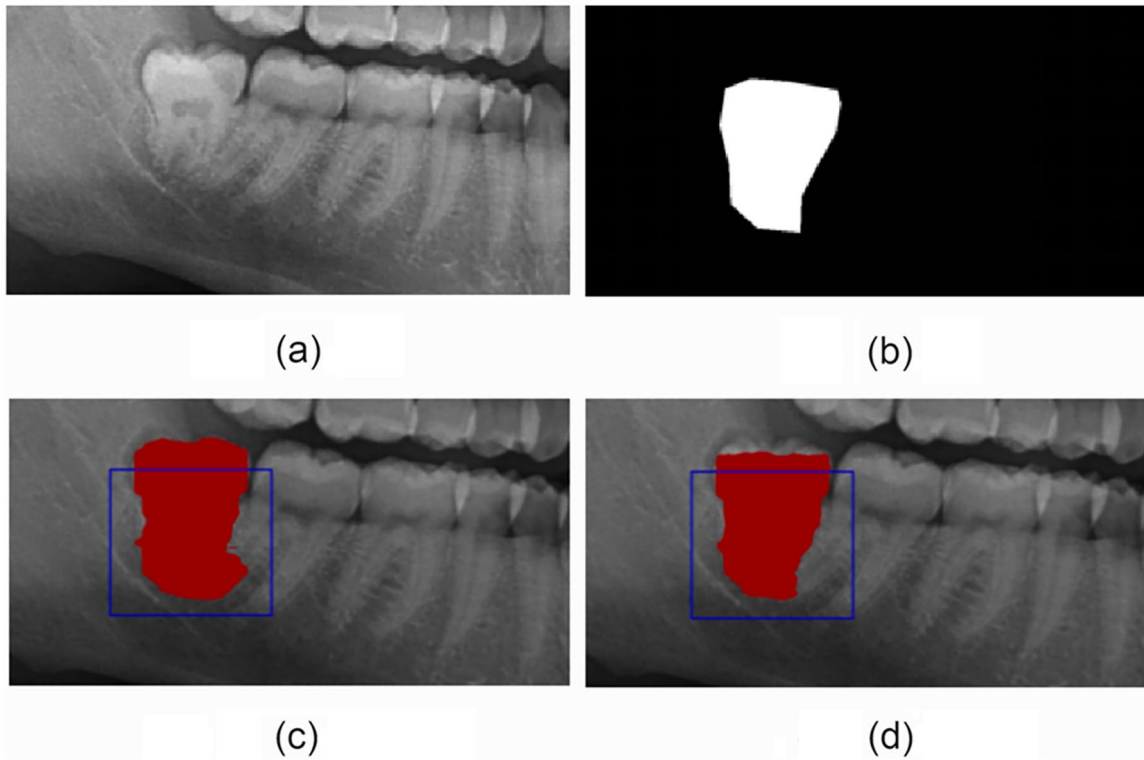**Figure 3.** The line and radar chart of the performance of SAM family.

Figure 4. The illustrartion of the segmentation performance of SAM family: (a) image, (b) mask, (c) SAM (huge), and (d) SAM (large).

ORCID iD

He Zhicheng https://orcid.org/0009-0000-6233-440X

REFERENCES

1. Ness G, Blakey G, Hechler B. *Impacted teeth. Peterson's principles of oral and maxillofacial surgery*. Springer. 2022:131-169.
2. Orhan K, Bilgir E, Bayrakdar IS, et al. Evaluation of artificial intelligence for detecting impacted third molars on cone-beam computed tomography scans. *J Stomatol Oral Maxillofac Surg*. 2021;122:333-337.
3. Faure J, Engelbrecht A. Impacted tooth detection in panoramic radiographs. In: *Advances in Computational Intelligence. 16th International Work-Conference on Artificial Neural Networks, IWANN 2021, Virtual Event*; 16–18 June, 2021. Springer International Publishing; 2021:525-536.
4. Giger ML. Computer-aided diagnosis in radiology. *Acad Radiol*. 2002;9:1-3.
5. Tajima S, Okamoto Y, Kobayashi T, et al. Development of an automatic detection model using artificial intelligence for the detection of cyst-like radiolucent lesions of the jaws on panoramic radiographs with small training datasets. *J Oral Maxillofac Surg Med Pathol*. 2022;34:553-560.
6. Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Comput Med Imaging Graph*. 2007;31:198-211.
7. Shin HC, Roth HR, Gao M, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging*. 2016;35:1285-1298.
8. Liu L, Ouyang W, Wang X, et al. Deep learning for generic object detection: a survey. *Int J Comput Vis*. 2020;128:261-318.
9. Zhu H, Cao Z, Lian L, et al. CariesNet: a deep learning approach for segmentation of multi-stage caries lesion from oral panoramic X-ray image. *Neural Comput Appl*. 2022;35:1-9.
10. Majanga V, Viriri S. Automatic blob detection for dental caries. *Appl Sci*. 2021;11:9232.

11. Hossen R, Arefin M, Nasir Uddin M. Object detection on dental X-ray images using region-based convolutional neural networks. In: *Machine Intelligence and Data Science Applications: Proceedings of MIDAS 2021*. Springer Nature Singapore; 2022:341-353.

12. Lin X, Hong D, Zhang D, Huang M, Yu H. Detecting proximal caries on periapical radiographs using convolutional neural networks with different training strategies on small datasets. *Diagnostics*. 2022;12:1047.

13. Zhu X, Wang F, Xie Y, et al. Periodontitis auxiliary diagnosis based on deep learning and oral dental x-ray images. *Russ J Nondestruct Test*. 2023;59: 487-500.

14. Bayrakdar IS, Orhan K, Çelik, et al. A U-net approach to apical lesion segmentation on panoramic radiographs. *Biomed Res Int*. 2022;2022: 7035367.

15. Shon HS, Kong V, Park JS, et al. Deep learning model for classifying periodontitis stages on dental panoramic radiography. *Appl Sci*. 2022;12: 8500.

16. Jiang L, Chen D, Cao Z, et al. A two-stage deep learning architecture for radiographic staging of periodontal bone loss. *BMC Oral Health*. 2022; 22:106.

17. Okazaki S, Mine Y, Iwamoto Y, et al. Analysis of the feasibility of using deep learning for multiclass classification of dental anomalies on panoramic radiographs. *Dent Mater J*. 2022;41:889-895.

18. Başaran M, Çelik Ö, Bayrakdar IS, et al. Diagnostic charting of panoramic radiography using deep-learning artificial intelligence system. *Oral Radiol*. 2022;38:363-369.

19. Guo J, Tsai PW, Xue X, et al. TVGG dental implant identification system. *Front Pharmacol*. 2022;13:948283.

20. Sukegawa S, Yoshii K, Hara T, et al. Multi-task deep learning model for classification of dental implant brand and treatment stage using dental panoramic radiograph images. *Biomolecules*. 2021;11:815.

21. Imak A, Çelebi A, Polat O, Türkoğlu M, Şengür A. ResMIBCU-Net: an encoder-decoder network with residual blocks, modified inverted residual block, and bi-directional ConvLSTM for impacted tooth segmentation in panoramic X-ray images. *Oral Radiol*. 2023;39:614-628.

22. Kuwada C, Ariji Y, Fukuda M, et al. Deep learning systems for detecting and classifying the presence of impacted supernumerary teeth in the maxillary incisor region on panoramic radiographs. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2020;130:464-469.

23. Zhang W, Li J, Li ZB, Li Z. Predicting postoperative facial swelling following impacted mandibular third molars extraction by using artificial neural networks evaluation. *Sci Rep*. 2018;8:12281.

24. Celik ME. Deep learning based detection tool for impacted mandibular third molar teeth. *Diagnostics*. 2022;12:942.

25. Kirillov A, Mintun E, Ravi N, et al. Segment anything. arXiv preprint arXiv 2304.02643. 2023.

26. Hesamian MH, Jia W, He X, Kennedy P. Deep learning techniques for medical image segmentation: achievements and challenges. *J Digit Imaging*. 2019;32:582-596.

27. Cheng J, Ye J, Deng Z, et al. Sam-med2d. ArXiv preprint arXiv 2308.16184. 2023.

28. Ma J, He Y, Li F, et al. Segment anything in medical images. *Nat Commun*. 2024;15:654.

29. Hamamci IE, Er S, Simsar E, Yuksel, et al. Dentex: An abnormal tooth detection with dental enumeration and diagnosis benchmark for panoramic x-rays. arXiv preprint arXiv 2305.19112. 2023.